Measuring Overuse With Electronic Health Records Data

Thomas Isaac, MD, MBA, MPH; Meredith B. Rosenthal, PhD; Carrie H. Colla, PhD; Nancy E. Morden, MD, MPH; Alexander J. Mainor, JD, MPH; Zhonghe Li, MS; Kevin H. Nguyen, MS; Elizabeth A. Kinsella, BA; and Thomas D. Sequist, MD, MPH

ABSTRACT

OBJECTIVES: To measure overuse of low-value care using electronic health record (EHR) data and manual chart review and to evaluate whether certain low-value services are better captured using EHR data.

STUDY DESIGN: We implemented algorithms to extract performance on 13 Choosing Wisely–identified healthcare services using EHR data at a large physician practice group between 2011 and 2013.

METHODS: We calculated rates of overuse using automated EHR extracts. We manually reviewed the charts for 200 cases of overuse for each measure to determine if they had clinical risk factors that could explain use of the low-value service and then calculated adjusted rates of overuse. We explored trends in overuse for each low-value service in the 3-year duration using logistic regression.

RESULTS: Unadjusted rates of overuse ranged from 0.2% to 92%. Automated EHR extracts and manual chart review identified explanatory risk factors for most measures, although the magnitude varied: for some measures (eg, bone densitometry exam for women younger than 65 years), manual chart review did not identify many additional risks (3.0%). In contrast, in patients who had sinus computed tomography or an antibiotic prescription for uncomplicated acute rhinosinusitis, manual chart review identified more explanatory risk factors (22.5%) than the automated EHR extract (9.5%). Adjusted rates of overuse ranged from 0.2% to 61.9%. Eight services demonstrated a statistically significant decrease in overuse over 3 years, while 1 increased significantly.

CONCLUSIONS: The use of EHR data, both extracted and manually abstracted, provides an opportunity to more accurately and reliably identify overuse of low-value healthcare services.

Am J Manag Care. 2018;24(1):19-25

In recent years, various stakeholders have sought strategies to reduce healthcare costs while improving the quality of care.[1] One such strategy is to define, identify, and reduce wasteful spending on the delivery of unnecessary healthcare services. In 2012, the American Board of Internal Medicine launched the Choosing Wisely campaign, an initiative aimed at engaging patients and physicians in discussions regarding how to avoid unnecessary healthcare.[2] More than 80 specialty societies have developed a list of recommendations to help physicians avoid unnecessary treatments, which include imaging studies, surgical procedures, medications, and laboratory testing.[3]

Findings of some early studies suggest that the Choosing Wisely initiative is showing modest results in its first years[4,5]; however, the success of this initiative, and others like it, is critically dependent on our ability to accurately measure healthcare overuse.[6,7] Most studies in the literature have relied on administrative claims data,[8-11] although these data are better suited for measuring simple processes of care and are not suitable to capture the nuances of appropriateness of care. Studies estimating overuse with administrative claims report that claims data can lack clinical context and that the use of such data risks misclassification of potentially appropriate testing or imaging.[12-14] Electronic health records (EHRs) may capture more detailed clinical information and context behind inappropriate use that are not available in claims data.[15-20]

The goal of our study was to measure overuse defined by the Choosing Wisely initiative using a combination of structured data extracts and manual chart review from EHRs. Using this combination of rich clinical information, we sought to determine whether certain types of low-value healthcare services are better captured than others using EHR data.

METHODS

Study Setting

We conducted this study at Atrius Health, a large physician practice group with 900 physicians and more than 400 advanced practice

TAKEAWAY POINTS

- In 13 low-value tests and screenings, we found varying levels of overuse using both automatically extracted electronic health record (EHR) data and manual chart review.
- Although several studies reporting overuse of low-value health services rely on administrative claims data, we found that EHR data may accurately and reliably measure overuse.
- EHR data can provide important insights on the presence of clinical risk factors that may trigger or explain the use of low-value health services.
- EHR data extracts and manual chart review should be considered alongside other methodologies of measuring overuse to develop higher-value local treatment norms for clinicians.

clinicians that provides primary and specialty care for more than 740,000 patients in eastern Massachusetts. Atrius Health physicians have utilized an integrated EHR (Epic Systems) since 1995 to support computerized outpatient ordering of medications, laboratory tests, and radiologic studies. All outpatient encounters are entered into the medical record, including vital signs, clinical notes, diagnostic and procedure codes, and all laboratory and radiology results.

Choosing Wisely Recommendations

A team of 2 physician health services researchers and 2 health economists reviewed available Choosing Wisely recommendations for inclusion in the study. We selected 13 Choosing Wisely recommendations (**Table 1**) for analysis based on their relevance to both primary and specialty care; their inclusion of medications, procedures, and laboratory testing; and their focus on ambulatory care, given that Atrius Health does not provide hospital care and therefore does not capture these data reliably.

We implemented algorithms to electronically capture performance on the selected Choosing Wisely recommendations between 2011 and 2013 using automated data extracts from coded data in the EHR. These data included the problem and medication lists and all diagnostic, lab, and procedure codes entered by clinicians during clinical encounters (**eAppendix I** [eAppendices available at **ajmc.com**]). The extracts implemented numerator and denominator inclusion definitions and captured any potential denominator exclusion criteria for each measure. We also used manual chart reviews to collect information on "explanatory risk factors" from the EHR, which were conditions that indicated why the Choosing Wisely test or treatment would have been clinically appropriate. Two measures (repeat bone densitometry [DEXA] exams and repeat endoscopy for Barrett's esophagus) were assessed in years prior to 2011, as they required historic data to measure whether the repeat exams occurred between 2011 and 2013.

We measured overuse via a denominator at either the exam or patient level, based on whether we could identify a patient population accurately. Measures that defined the patient denominator based only on demographic criteria and required little clinical evaluation (eg, cancer screening measures for women) were measured using an exam-level denominator. In contrast, measures that defined the patient denominator based on presentation to the office with a symptom (eg, headache or syncope) were measured using a patient-level denominator. This approach allowed us to avoid making assumptions about whether patients were receiving their care at the physician group practice during the entire study period and should thus be included in the denominator.

Manual Chart Review

From the electronic data extracts, we selected a random sample of 200 cases of overuse for each measure based on its numerator definition (eAppendix I) from each study year for manual chart review. Manual chart review data were considered the gold standard for our analyses, as such data are generally the most robust clinical information source available. The chart review team consisted of 2 board-certified internal medicine physicians and a research assistant. The manual chart reviews focused on the index clinical note associated with ordering of the targeted overuse service, as well as any other information referenced in this index clinical note, such as prior clinical notes, study results, or medication lists.

The review team collected information on whether the EHR data extracts were working as intended in capturing the numerator, denominator, and exclusion definitions of each measure; developed a list of clinical explanatory risk factors for each measure based on guidelines; and examined whether these were present in the clinical notes or test ordering documentation.

Data Analysis

We fit multivariable logistic regression models to assess trends of overuse using EHR data for each Choosing Wisely recommendation, incorporating definitions of overuse (eAppendix I) and controlling for each study year between 2009 and 2013.

The chart review team collected data on the number of explanatory factors found from EHR extracts and manual chart review. In cases where there was confusion about whether a test or procedure might have been clinically indicated, the 3 reviewers used a consensus approach to determine whether a legitimate explanatory risk factor was present.

We reported unadjusted and adjusted performance rates after excluding patients with explanatory risk factors. All analyses were conducted using Stata version 14 (StataCorp LP; College Station, Texas). The study protocol was approved by the Partners Healthcare System Human Studies Review Committee.

RESULTS

The prevalence of overuse varied widely, both among those measures with exam-level denominators and those with patient-level

TABLE 1. Select Choosing Wisely Recommendations and Measure Specifications

Choosing Wisely Recommendation	Performance Measure	Denominator[a]
Don't use DEXA screening for osteoporosis in women younger than 65 years without risk factors	% of all DEXA exams performed on women aged 18-64 years without risk factors	Exam-level
Don't routinely repeat DEXA scans more often than once every 2 years	% of DEXA exams on women 18 years or older followed by a repeat within 2 years	Exam-level
Don't perform cardiac testing for low-risk patients without symptoms	% of electrocardiograms, stress tests, and cardiac imaging exams performed on low-risk patients 18 years or older without symptoms	Exam-level
Don't perform Pap smears on women younger than 21 years	% of all Pap smears performed on women aged 18-20 years	Exam-level
Don't perform Pap smears on women older than 65 years with adequate prior screening	% of all Pap smears performed on women older than 65 years with adequate prior screening based on historic Pap smear results	Exam-level
Don't perform Pap smears on women with total hysterectomy for noncancer disease	% of all Pap smears performed on women 18 years or older with total hysterectomy for noncancer disease	Exam-level
Don't repeat endoscopy within 3 years for Barrett's esophagus if prior exam is normal	% of normal endoscopies for patients 18 years or older with Barrett's esophagus followed by a repeat endoscopy within 3 years	Exam-level
Don't perform population-based screening for 25-OH-vitamin D deficiency	% of 25-OH-vitamin D tests performed on patients 18 years or older with no medical indication	Exam-level
Don't obtain brain imaging for simple syncope	% of patients 18 years or older presenting with syncope and no alarm features who have either head CT or MRI performed within 1 month	Patient-level
Don't do imaging for low back pain within first 6 weeks	% of patients 18 years or older presenting with low back pain and no alarm symptoms who have back x-ray, CT, or MRI performed within 6 weeks of presentation	Patient-level
Don't use opioid or butalbital treatment for migraine, except as a last resort	% of patients 18 years or older presenting with migraine who received opioid or butalbital treatment	Patient-level
Don't order sinus CT or prescribe antibiotics for uncomplicated acute rhinosinusitis	% of patients 18 years or older presenting with uncomplicated acute rhinosinusitis who had a sinus CT within 1 month or were prescribed antibiotics within 3 weeks	Patient-level
Don't do imaging for uncomplicated headache	% of patients 18 years or older presenting with headache and no alarm features who receive a head MRI or CT within 1 month	Patient-level

CT indicates computed tomography; DEXA, bone densitometry; MRI, magnetic resonance imaging; OH, hydroxy.
[a]Exam-level measures have 1 record per exam in denominator population (ie, patients repeated if more than 1 exam is present during measurement period). Patient-level measures have 1 record per patient in denominator population.

denominators, after accounting for explanatory risk factors found in the EHR (**Table 2**). In 2013, among exam-level measures, the prevalence of overuse ranged from a low of 0.2% (proportion of Pap smears performed on women aged 18-21 years) to a high of 57% (proportion of DEXA exams performed on women aged 18-65 years). Among patient-level measures in 2013, the prevalence of overuse ranged from a low of 8% (use of head imaging for syncope) to a high of 92% (use of sinus computed tomography [CT] or antibiotics for acute rhinosinusitis).

Among the 12 measures with 3 years of data, the prevalence of overuse demonstrated a statistically significant decrease over time for 8 (67%) measures. There was a significant increase in prevalence for only 1 (8%) measure, the use of head imaging for patients with uncomplicated headache.

We found wide variation in the number of risk factors identified by the automated EHR extract across measures (**Table 3**). In Vitamin D deficiency screening, imaging for low-back pain, and both DEXA measures, over half of the sample was shown to have risk factors that could explain the test or screening using the automated EHR extract; however, other measures, like Pap smears performed on women younger than 21 years, found no explanatory risk factors when using the EHR extract alone.

Across nearly all measures, the manual chart review identified additional explanatory risk factors, although the magnitude varied by measure (Table 3). Manual chart review identified few additional explanatory risk factors in DEXA exams performed on women younger than 65 years (3.0%) and prescription of opioids or butalbital treatment for patients with migraine (5.0%), but more in patients who had a sinus CT or antibiotic prescription for uncomplicated acute rhinosinusitis (22.5%) and Pap smears performed on women with total hysterectomy for noncancer disease (29.0%). **eAppendix II** presents the explanatory risk factors for each measure.

We calculated an adjusted overuse rate in 2013 based on the proportion of EHR-identified overuse that manual chart review

TABLE 2. Trends in Overuse Based on Automated EHR Extracts[a]

Performance Measure	2009 n (%)	2010 n (%)	2011 n (%)	2012 n (%)	2013 n (%)	*P*
% of all DEXA exams performed on women aged 18-64 years without risk factors	–	–	5865 (62)	5315 (60)	5229 (57)	<.001
% of DEXA exams on women 18 years or older followed by a repeat within 2 years	1523 (12)	1430 (11)	1357 (11)	–	–	.01
% of electrocardiograms, stress tests, and cardiac imaging exams performed on low-risk patients 18 years or older without symptoms	–	–	16,302 (19)	16,062 (18)	14,801 (16)	<.001
% of all Pap smears performed on women aged 18-20 years	–	–	851 (1.1)	360 (0.5)	129 (0.2)	<.001
% of all Pap smears performed on women older than 65 years with adequate prior screening based on historic Pap smear results	–	–	30,701 (37)	25,302 (35)	20,256 (32)	<.001
% of all Pap smears performed on women 18 years or older with total hysterectomy for noncancer disease	–	–	656 (0.8)	535 (0.7)	433 (0.7)	.04
% of normal endoscopies for patients 18 years or older with Barrett's esophagus followed by a repeat endoscopy within 3 years[b]	–	413 (70)	–	–	–	–
% of 25-OH-vitamin D tests performed on patients 18 years or older with no medical indication	–	–	29,539 (51)	20,183 (40)	19,332 (38)	<.001
% of patients 18 years or older presenting with syncope and no alarm features who have either head CT or MRI performed within 1 month	–	–	216 (7)	229 (8)	242 (8)	.14
% of patients 18 years or older presenting with low back pain and no alarm symptoms who have back x-ray, CT, or MRI performed within 6 weeks	–	–	2650 (18)	2116 (16)	1969 (16)	<.001
% of patients 18 years or older presenting with migraine who received opioid or butalbital treatment	–	–	1623 (17)	1138 (18)	1070 (18)	.08
% of patients 18 years or older presenting with uncomplicated acute rhinosinusitis who had:						
Sinus CT within 1 month	–	–	35 (0.8)	28 (0.7)	17 (0.5)	.09
Antibiotics within 3 weeks	–	–	3883 (93)	3501 (93)	3104 (91)	.07
Either sinus CT or antibiotics	–	–	3888 (93)	3503 (93)	3107 (91)	.06
% of patients 18 years or older presenting with headache and no alarm features who received a head MRI or CT within 1 month	–	–	1694 (18)	1582 (20)	1573 (20)	<.001

CT indicates computed tomography; DEXA, bone densitometry; EHR, electronic health record; MRI, magnetic resonance imaging; OH, hydroxy.
[a]Multivariable logistic regression models were used to calculate trends over time using each performance measure as dependent variable, based on definitions of overuse in eAppendix I, controlling for each study year between 2009 and 2013.
[b]Trend data were not available for the repeat endoscopy measure because we identified patients with Barrett's esophagus in 2010 and followed them for 3 years.

determined had risk factors that could explain the test or screening (Table 3). The prevalence of overuse after adjustment ranged between 0.2% for Pap smears on women younger than 21 years and 61.9% for patients with uncomplicated acute rhinosinusitis who had a sinus CT or were prescribed antibiotics.

DISCUSSION

There is widespread agreement that low-value services are a substantial problem leading to enormous waste in healthcare, and Choosing Wisely recommendations are often suggested as a tool to begin to define and identify waste. We examined how easily EHR data from a large ambulatory care center could be used to identify overuse on a selection of Choosing Wisely measures, and we found mixed results.

Previous study findings indicate that administrative claims are a useful, albeit limited, source of data for identifying overuse.[14] Primarily used for documenting diagnoses and procedures provided for the purpose of payment, administrative claims data do not provide detailed clinical context, which can ultimately misclassify, underestimate, or overestimate indicated tests or screenings.[21] Reliance on claims alone may misclassify a clinically appropriate test or screening as overuse of low-value care, as patient history is an integral factor in clinical decision making.[22] Several studies, including our own, cite reliance solely on administrative claims data for measuring overuse as a limitation for accurately reporting overuse.[15,23,24]

TABLE 3. Explanatory Risk Factors and Overuse Identified Through EHR and Manual Chart Review[a]

Choosing Wisely Recommendation	N = 200 per Recommendation: Explanatory Risk Factors Identified in Automated EHR Extract n (%)	Additional Explanatory Risk Factors Identified in Manual Chart Review n (%)	Sample Remaining After EHR and Chart Review Risk Factor Exclusions n (%)	Unadjusted Overuse Rate, 2013 (%)[b]	Adjusted Overuse Rate After Chart Review (%)[c]
% of DEXA exams performed on women younger than 65 years without risk factors	129 (64.5%)	6 (3.0%)	65 (32.5%)	57.0	18.5
% of DEXA exams followed by a repeat within 2 years	155 (77.5%)	42 (24.0%)	3 (1.5%)	–	N/A
% of cardiac testing performed on low-risk patients without symptoms	0 (0%)	37 (18.5%)	163 (81.5%)	16.1	13.0
% of Pap smears performed on women aged 18-20 years	0 (0%)	22 (11.0%)	178 (89.0%)	0.2	0.2
% of Pap smears performed on women older than 65 years with adequate prior screening	27 (13.5%)	31 (15.5%)	142 (71.0%)	32.4	22.7
% of Pap smears performed on women with total hysterectomy for noncancer disease	0 (0%)	58 (29.0%)	142 (71.0%)	0.7	0.5
% of normal endoscopies for patients with Barrett's esophagus followed by a repeat within 3 years	44 (22.0%)	29 (14.5%)	127 (36.5%)	–	N/A
% of 25-OH-vitamin D deficiency tests performed for nonindicated reasons	106 (53.0%)	18 (9.0%)	76 (38.0%)	38.0	14.4
% of patients with simple syncope who had brain imaging	61 (30.5%)	72 (36.0%)	67 (33.5%)	8.2	2.7
% of patients with imaging performed for low back pain within first 6 weeks	110 (55.0%)	24 (12.0%)	66 (33.0%)	16.0	5.3
% of patients with migraine who received opioid or butalbital treatment	25 (12.5%)	10 (5.0%)	165 (82.5%)	18.4	14.9
% of patients with uncomplicated acute rhinosinusitis who had sinus CT or were prescribed antibiotics	19 (9.5%)	45 (22.5%)	136 (68.0%)	91.4	61.9
CT (n = 3)	1 (33.3%)	0 (0.0%)	2 (66.7%)	0.5	0.3
Antibiotics (n = 199)	19 (9.5%)	45 (22.5%)	135 (67.8%)	90.9	61.7
% of patients with uncomplicated headache who received imaging	47 (23.5%)	29 (14.5%)	124 (62.0%)	20.2	12.4

CT indicates computed tomography; DEXA, bone densitometry; EHR, electronic health record; N/A, not applicable; OH, hydroxy.
[a]For each Choosing Wisely measure, a random sample of 200 cases were selected that met initial criteria for overuse based on EHR data. We then used EHR data extracts and manual chart reviews to find additional explanatory risk factors that could explain the testing or treatment.
[b]Unadjusted overuse rate reflects the estimated overuse based on automated EHR extracts in 2013, which accounts for risk factors identified in structured EHR data fields that were determined to be legitimate exclusions.
[c]Adjusted overuse rate after chart review reflects the proportion of EHR-identified overuse in 2013 (see Table 2) that accounts for risk factors identified in manual chart review that were determined to be legitimate exclusions.

Our findings suggest that EHR data can be an important, although variable, source of information in identifying overuse of clinical services. For some measures of overuse, such as Pap smears in women younger than 21 years, structured EHR extracts were sufficient for identifying rates of overuse and relevant risk factors. In these cases, manual chart review added little insight into the potential clinical justification for a test or screening.

For most other measures, the combination of EHR data and manual chart review provided valuable information and elucidated some of the inherent complexities in overuse measures. For instance, cases of DEXAs in women younger than 65 years were

easy to identify in the EHR, although clinical risk factors, such as previous osteopenia, were often identified in manual chart review. The initial diagnosis of osteopenia was often obtained from a DEXA that was not clearly indicated, thereby suggesting that an initial low-value test or screening may lead to subsequent low-value services. Further, there is debate on what the correct duration of follow-up should be after identifying mild or moderate osteopenia.[25] Understanding these clinical nuances that may explain imaging, testing, or procedures is important, given the potential implication on costs (eg, the estimated cost of a single DEXA exam is approximately $125).[26] The measure regarding overuse of antibiotics for sinusitis also proved challenging. Most cases of sinusitis identified using the EHR received antibiotics, which might represent a coding bias of clinicians in which the diagnosis is only listed when treated.

Although there is variation in magnitude across measures, our study results suggest that EHR data provide important insights on overuse and presence of risk factors for several Choosing Wisely recommendations. However, both claims data and EHR data have limitations that can overestimate overuse, as the presence of risk factors is often only captured in chart review. We used manual chart review as the gold standard in our study, but recognize the labor-intensive nature of this methodology. Development of more automated text-based extracts (eg, natural language processing) could provide a less resource-intensive means to identify legitimate explanatory clinical risk factors of overuse. Further, as practices incorporate clinical decision support to identify low-value testing in real time and query providers to specify a clinical justification, the utility of EHR data extracts should improve.

Limitations

Our analysis has a number of limitations. We examined the use of EHR data in a large ambulatory care system that uses Epic software. The data warehouse structure and information available for procedures, medications, and laboratory tests likely have large variations compared with other EHRs. Second, we examined only a selection of Choosing Wisely recommendations, although the sample had variety in measures pertaining to medications, imaging, and procedures. Third, our study relied on manual chart review as the gold standard for determining overuse. Although manual chart review provides more clinical information than administrative claims data, we relied on information documented in the patient chart and therefore may be missing data that were not documented. Fourth, our chart reviews only examined clinical information from the encounter associated with the test order of each Choosing Wisely measure. A more thorough chart review looking back at previous notes and outside notes would likely yield more explanatory information, although this type of review requires more resources to perform. Finally, our EHR extracts and chart reviews examining explanatory factors and risk factors for each measure are open to clinical interpretation, and the clinical opinion of reviewers would impact the reproducibility of our results.

CONCLUSIONS

As clinicians and policy makers continue to gather data on overuse of low-value services, the methodologies and data sources utilized to measure overuse have become increasingly important. Developing more accurate and reliable calculations of overuse would be instrumental for policy makers and providers to identify opportunities for changing care delivery. Our work suggests that EHRs are an important source of data to quantify overuse and that EHRs can capture clinical information that often explains why a test or treatment is clinically indicated. Further, manual chart review, although more resource-intensive, may identify the presence of important risk factors that automated EHR data extracts cannot, and it should be considered alongside other methodologies of measuring overuse. The data from such manual chart reviews might be particularly important when engaging clinicians in the development and implementation of care delivery practices that reduce overuse of low-value services. ■

Author Affiliations: Atrius Health (TI), Newton, MA; Department of Health Policy and Management, Harvard T.H. Chan School of Public Health (MBR, ZL, KHN), Boston, MA; The Dartmouth Institute for Health Policy and Clinical Practice (CHC, NEM, AJM), Lebanon, NH; Norris Cotton Cancer Center, Dartmouth-Hitchcock Medical Center (CHC), Lebanon, NH; Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth (NEM), Lebanon, NH; Division of General Internal Medicine, Brigham and Women's Hospital (EAK, TDS), Boston, MA; Partners HealthCare (EAK, TDS), Boston, MA; Department of Health Care Policy, Harvard Medical School (TI, TDS), Boston, MA.

Source of Funding: The Commonwealth Fund.

Author Disclosures: The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (TI, MBR, CHC, NEM, EAK, TDS); acquisition of data (TI, MBR, ZL, EAK, TDS); analysis and interpretation of data (TI, MBR, CHC, NEM, AJM, ZL, KHN, EAK, TDS); drafting of the manuscript (TI, MBR, CHC, NEM, AJM, KHN); critical revision of the manuscript for important intellectual content (TI, MBR, CHC, NEM, AJM, TDS); statistical analysis (MBR, CHC, ZL, KHN, TDS); obtaining funding (MBR, CHC, TDS); administrative, technical, or logistic support (CHC, AJM, KHN, EAK); and supervision (MBR).

Address Correspondence to: Thomas Isaac, MD, Atrius Health, 275 Grove St, Ste 3-100, Auburndale, MA 02466. Email: thomas_isaac@atriushealth.org.

REFERENCES

1. Schwartz AL, Chernew ME, Landon BE, McWilliams JM. Changes in low-value services in year 1 of the Medicare Pioneer Accountable Care Organization program. *JAMA Intern Med.* 2015;175(11):1815-1825. doi: 10.1001/jamainternmed.2015.4525.
2. About. Choosing Wisely website. choosingwisely.org/about-us. Accessed December 11, 2015.
3. Lists. Choosing Wisely website. choosingwisely.org/doctor-patient-lists/. Accessed December 11, 2015.
4. Rosenberg A, Agiro A, Gottlieb M, et al. Early trends among seven recommendations from the Choosing Wisely campaign. *JAMA Intern Med.* 2015;175(12):1913-1920. doi: 10.1001/jamainternmed.2015.5441.
5. Grover M, McLemore R, Tilburt J. Clinicians report difficulty limiting low-value services in daily practice. *J Prim Care Community Health.* 2016;7(2):135-138. doi: 10.1177/2150131915624112.
6. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care*. 1986;2(1):53-63.

7. Smith M, Saunders R, Stuckhardt L, McGinnis JM, eds. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America.* Washington, DC: National Academies Press; 2013. nap.edu/catalog/13444/best-care-at-lower-cost-the-path-to-continuously-learning.
8. Colla CH. Swimming against the current—what might work to reduce low-value care? *N Engl J Med.* 2014;371(14):1280-1283. doi: 10.1056/NEJMp1404503.
9. Schwartz AL, Landon BE, Elshaug AG, Chernew ME, McWilliams JM. Measuring low-value care in Medicare. *JAMA Intern Med.* 2014;174(7):1067-1076. doi: 10.1001/jamainternmed.2014.1541.
10. Reid RO, Rabideau B, Sood N. Low-value health care services in a commercially insured population. *JAMA Intern Med.* 2016;176(10):1567-1571. doi: 10.1001/jamainternmed.2016.5031.
11. Charlesworth CJ, Meath TH, Schwartz AL, McConnell KJ. Comparison of low-value care in Medicaid vs commercially insured populations. *JAMA Intern Med.* 2016;176(7):998-1004. doi: 10.1001/jamainternmed.2016.2086.
12. Bhatia RS, Levinson W, Shortt S, et al. Measuring the effect of Choosing Wisely: an integrated framework to assess campaign impact on low-value care. *BMJ Qual Saf.* 2015;24(8):523-531. doi: 10.1136/bmjqs-2015-004070.
13. Hong AS, Ross-Degnan D, Zhang F, Wharam JF. Small decline in low-value back imaging associated with the 'Choosing Wisely' campaign, 2012-14. *Health Aff (Millwood).* 2017;36(4):671-679. doi: 10.1377/hlthaff.2016.1263.
14. Tang PC, Ralston M, Arrigotti MF, Qureshi L, Graham J. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *J Am Med Inform Assoc.* 2007;14(1):10-15. doi: 10.1197/jamia.M2198.
15. Elshaug AG, McWilliams JM, Landon BE. The value of low-value lists. *JAMA.* 2013;309(8):775-776. doi: 10.1001/jama.2013.828.
16. Bailey SR, Heintzman JD, Marino M, et al. Measuring preventive care delivery: comparing rates across three data sources. *Am J Prev Med.* 2016;51(5):752-761. doi: 10.1016/j.amepre.2016.07.004.
17. Naessens JM, Ruud KL, Tulledge-Scheitel SM, Stroebel RJ, Cabanela RL. Comparison of provider claims data versus medical records review for assessing provision of adult preventive services. *J Ambul Care Manage.* 2008;31(2):178-186. doi: 10.1097/01.JAC.0000314708.65289.3b.
18. Baker DW, Qaseem A, Reynolds PP, Gardner LA, Schneider EC; American College of Physicians Performance Measurement Committee. Design and use of performance measures to decrease low-value services and achieve cost-conscious care. *Ann Intern Med.* 2013;158(1):55-59. doi: 10.7326/0003-4819-158-1-201301010-00560.
19. Heintzman J, Bailey SR, Hoopes MJ, et al. Agreement of Medicaid claims and electronic health records for assessing preventive care quality among adults. *J Am Med Inform Assoc.* 2014;21(4):720-724. doi: 10.1136/amiajnl-2013-002333.
20. Kern LM, Malhotra S, Barrón Y, et al. Accuracy of electronically reported "meaningful use" clinical quality measures: a cross-sectional study. *Ann Intern Med.* 2013;158(2):77-83. doi: 10.7326/0003-4819-158-2-201301150-00001.
21. Schlemmer E, Mitchiner JC, Brown M, Wasilevich E. Imaging during low back pain ED visits: a claims-based descriptive analysis. *Am J Emerg Med.* 2015;33(3):414-418. doi: 10.1016/j.ajem.2014.12.060.
22. Colla CH, Morden NE, Sequist TD, Schpero WL, Rosenthal MB. Choosing wisely: prevalence and correlates of low-value health care services in the United States [erratum in *J Gen Intern Med.* 2016;31(4):450. doi: 10.1007/s11606-015-3420-5]. *J Gen Intern Med.* 2015;30(2):221-228. doi: 10.1007/s11606-014-3070-z.
23. Colla CH, Sequist TD, Rosenthal MB, Schpero WL, Gottlieb DJ, Morden NE. Use of non-indicated cardiac testing in low-risk patients: Choosing Wisely. *BMJ Qual Saf.* 2015;24(2):149-153. doi: 10.1136/bmjqs-2014-003087.
24. Backhus LM, Farjah F, Varghese TK, et al. Appropriateness of imaging for lung cancer staging in a national cohort. *J Clin Oncol.* 2014;32(30):3428-3435. doi: 10.1200/JCO.2014.55.6589.
25. Gourlay ML, Fine JP, Preisser JS, et al; Study of Osteoporotic Fractures Research Group. Bone-density testing interval and transition to osteoporosis in older women. *N Engl J Med.* 2012;366(3):225-233. doi: 10.1056/NEJMoa1107142.
26. Bone-density tests. Choosing Wisely website. choosingwisely.org/patient-resources/bone-density-tests. Published May 2012. Accessed December 11, 2015.

Full text and PDF at **www.ajmc.com**

eAppendix I. Electronic Health Record Reported Measure Data Specifications – Numerator and Denominator Inclusion Definitions

Choosing Wisely Recommendation	EHR Reported Measure	Numerator	Denominator[a]
Don't use DEXA screening for osteoporosis in women younger than 65 without risk factors	% of DEXA exams performed on women younger than 65 without risk factors	DEXA exams performed on 18<=women<65	DEXA exams on women >=18
Don't routinely repeat DEXA scans more often than once every two years	% of DEXA exams followed by a repeat within 2 years	DEXA exams on women > 18 repeated within 2 years of prior exam	DEXA exams on women >=18
Don't perform Pap smears on women younger than 21	% of Pap smears performed on 18<=women<21	Pap smears on 18<=women<21	Pap smears on women >=18
Don't perform Pap smears on women >65 with adequate prior screening	% of Pap smears performed on women >65 with adequate prior screening	Pap smears on women >65 and adequate prior screening	Pap smears on women >=18
Don't perform Pap smear test on women with total hysterectomy for non-cancer disease	% of Pap smears performed on women with total hysterectomy for non-cancer disease	Pap smears on patients with total hysterectomy for non-cancer disease	Pap smears on women >=18
Don't repeat endoscopy within 3 years for Barrett's esophagus if prior exam is normal	% of normal endoscopies for Barrett's patients followed by a repeat within 3 years	Repeat endoscopies within 3 years of prior normal endoscopy for Barrett's	Normal-result endoscopies for Barrett's patients >=18
Don't perform population based screening for 25-OH-Vitamin D deficiency	% of 25-OH-Vitamin D deficiency screening performed for non-indicated reasons	25-OH-Vitamin D test for patients>=18 with no indication	25-OH-Vitamin D test for adults >= 18
Don't obtain brain imaging for simple syncope	% of patients with simple syncope who had brain imaging	Patients >=18 obtaining brain head CT or MRI/MRA	Patients >=18 with simple syncope
Don't do imaging for low back pain within first 6 weeks	% of patients with imaging performed for low back pain within first 6 weeks	Patients >=18 with low back pain and imaging	Patients >=18 with low back pain

		performed within first 6 weeks	
Don't use opioid or butalbital treatment for migraine, except as a last resort	% of patients with migraine who received opioid or butalbital treatment	Prescription for opioid or butalbital following evaluation for migraine	Patients >=18 with migraine headache
Don't order sinus CT or prescribe antibiotics for uncomplicated acute rhinosinusitis	% of patients with uncomplicated acute rhinosinutis who had a sinus CT or were prescribed antibiotics	Had sinus CT OR were prescribed antibiotics	Patients >= 18 with at least one face-to-face visit for rhinosinusitis
Don't do imaging for uncomplicated headache	% of patients with uncomplicated headache who receive imaging	Patients >=18 with brain imaging done for uncomplicated headache	Patients >=18 with uncomplicated headache

[a]Exam level measures have 1 record per exam in Denominator (ie, patients repeated if more than 1 exam). Patient measures have 1 record per patient identified in Denominator population.

eAppendix II. Electronic Health Record Reported Measure Data Specifications – Explanatory Risk Factors[a]

Measure	Risk Factors
Don't use DEXA screening for osteoporosis in women younger than 65 without risk factors	Y or N for each as of DEXA exam date: 1) Previous fracture - if Y, include dates of all fractures. 2) Smoking status 3) Oral or inhaled glucocorticoid use (medication order within 12 months prior to exam result date) 4) Anticonvulsant use (medication order within 12 months prior to exam result date) 5) Hx rheumatoid arthritis 6) Hx Osteoporosis 7) Osteopenia 8) Celiac disease 9) Chronic pancreatitis 10) Lactase deficiency 11) Crohn's disease 12) Ulcerative colitis 13) Vitamin D Deficiency 14) Hypogonadism 15) Hyperthyroidism 16) Menopause
Don't routinely repeat DEXA scans more often than once every 2 years	Y or N for each as of both (if applicable) DEXA exam dates: 1) Previous fracture - if Y, include dates of all fractures. 2) Smoking status 3) Oral or inhaled glucocorticoid use (medication order within 12 months prior to exam result date) 4) Anticonvulsant use (medication order within 12 months prior to exam result date) 5) Hx rheumatoid arthritis 6) Hx Osteoporosis 7) Hx Osteopenia 8) Celiac disease 9) Chronic pancreatitis 10) Lactase deficiency 11) Crohn's disease

Measure	Risk Factors
	12) Ulcerative colitis 13) Vitamin D Deficiency 14) Hypogonadism 15) Hyperthyroidism 16) Menopause
Don't perform cardiac testing for low risk patients without symptoms	As of cardiac test date: 1) Framingham Risk Score
Don't perform Pap smears on women younger than 21	As of Pap result date (Y/N): 1) History of cancer/pre-cancer (Y/N)?
Don't perform Pap smears on women >65 with adequate prior screening	As of Pap result date (Y/N): 1) History of cancer/pre-cancer (Y/N)?
Don't perform Pap smear test on women with total hysterectomy for non-cancer disease	N/A
Don't repeat endoscopy within 3 years for Barrett's esophagus if prior exam is normal	Code added AFTER 2010 endoscopy and BEFORE f/u endoscopy (Y/N): 1) Weight loss 2) Change in bowel habits 3) GI bleeding 4) Dysphagia 5) Odynophagia 6) Caustic ingestion 7) Esophageal varices 8) Portal hypertension 9) Removal of foreign body 10) Placement of feeding or drainage 11) Achalasia 12) Esophageal stricture

Measure	**Risk Factors**
	13) Esophageal cancer 14) Gastric cancer 15) Duodenal cancer 16) Pancreatic cancer 17) Gall Bladder cancer
Don't perform population based screening for 25-OH-Vitamin D deficiency	Appropriate indications include presence of the following prior to performance of the Vitamin D test, either on problem list or in face-to-face encounter diagnosis: 1) Kidney Disease 2) Osteoporosis 4) Osteopenia 5) Obesity 6) Parthyroidism 7) Liver disease 8) Anticonvulsant use (medication order within 12 months prior to Vit D test date 9) Celiac disease 10) Chronic pancreatitis 11) Lactase deficiency 12) Crohn's disease 13) Ulcerative colitis 14) Vitamin D Deficiency
Don't obtain brain imaging for simple syncope	Presence (Y/N) as of syncope diagnosis date: 1) Coronary heart disease 2) Epilepsy/ Seizure 3) Stroke/ TIA Presence (Y/N) of codes ***on the same day of visit:*** **1) Cardiac symptoms:** - Chest pain - Palpitations - Breathing difficulty (dyspnea, shortness of breath)

Measure	Risk Factors
	- arrhythmia **2) Neurologic symptoms:** - Visual changes - Seizures -Numbness/ Paresthesias - Hearing change - Change in smell - Incontinence - Aura - Headache
Don't do imaging for low back pain within first 6 weeks	Presence (Y/N) prior to INDEX diagnosis date OR between INDEX diagnosis date and imaging date, if applicable (if Y, include date of most recent face-to-face encounter dx): **** For each of the following 13 fields, please gather sub-fields A and B, where* *- A indicates "Y/N, within 12 months prior to INDEX dx date or btw dx and imaging date?" and* *- B indicates "Y/N, within 5 years prior to INDEX dx date or btw dx and imaging date?"* 1) Cancer (See **Cancer Codes** sheet) 1a) Cancer, Y/N, within 12 months prior to INDEX dx date or btw dx and imaging date? 1b) Cancer, Y/N, within 5 years prior to INDEX dx date or btw dx and imaging date? 2) Trauma 3) IV drug abuse 4) Neurologic impairment 5) Unexplained weight loss 6) Immunosuppression 7) **Oral glucocorticoid u**se (medication order within 12 months prior to exam result date) 8) Fever 9) Hx urinary infections 10) Hx abdominal aortic aneurysm 11) Low back pain dx present within 12 months prior to INDEX lbp dx **NOTE:** ***For this risk factor only,*** ***- A indicates "Y/N, within 12 months prior to INDEX dx date*** ***- B indicates "Y/N, within 5 years prior to INDEX dx date***

Measure	**Risk Factors**
	12) intraspinal abscess 13) Any "e" code
Don't use opioid or butalbital treatment for migraine, except as a last resort	Presence (Y/N) within 60 days prior to migraine diagnosis date: 1) External causes of injury: Any "E" code from ICD9 set (external causes) 2) Back pain 3) Fractures 4) All cancers, except non-melanoma skin cancers
Don't order sinus CT or prescribe antibiotics for uncomplicated acute rhinosinusitis	Identify presence of conditions that might indicated complicated or chronic rhinosinusitis (Y/N) diagnosed during the index encounter or within the following 1 month of the encounter: 1) Chronic sinusitis
Don't do imaging for uncomplicated headache	Within 5 years prior to index diagnosis: 1) Presence of cancer (Y/N), excluding non-melanoma skin cancer On Day of Diagnosis (Y/N): 1) Fever 2) Neurologic symptoms - Visual changes - Seizures - Numbness/Paresthesia - Hearing change - Change in smell - Incontinence - Aura 3) Syncope

[a]Two board certified internal medicine physicians and a research assistant collected information on whether we accurately captured the numerator, denominator, and exclusion definitions, and whether clinical explanatory risk factors were present.